1 .

Overview and Summary

Overview and Summary

INTRODUCTION

This paper is about the economic evaluation of diagnostic procedures. To be precise, it is about whether and when the benefits from a diagnostic procedure are worth its risks and costs.

The issue of economic evaluation is explored in the context of four common diagnostic X-ray procedures that together in 1970 accounted for almost half of all diagnostic X-ray examinations in the United States:

- the chest X-ray,
- the skull X-ray,
- the barium enema study, and
- the excretory urogram.

Since these are all long-established and widely used radiological procedures, evidence of their benefits, risks, and costs should be **as** comprehensive as that of most other diagnostic procedures. The methods used to evaluate these procedures should (and do) represent the general state-of-the-art in the assessment of most diagnostic technologies.

The four X-ray procedures are also interesting in their own right. As high-volume procedures, they use up substantial health care resources. Table 1 summarizes the findings of a 1970 survey of diagnostic X-ray procedure use in the United States. Of the 130 million diagnostic X-ray procedures performed in that year, 60 million were for the four procedures of interest here. The average amount billed by a sample of radiologists in California in 1975 for each of the four procedures is shown in table 2. Though billed charges overstate the amount actually received and data from California do not represent the nation, the total burden of expenditures for X-rays is substantial.

The appropriateness of use of each procedure has been the subject of intense debate within the

Table 2.—Average Amount Billed by Radiologists for Selected X-Ray Procedures, California Medicaid Program, 1975

Procedure	Average billed amount
Chest X-ray:	
Single	$13.03
Double	22.73
Complete	31.61
Barium enema:	
Colon, barium enema	45.32
With air contrast	64.29
Air contrast (independent procedure)	NA
Excretory urogram:	
Routine intravenous pyelography	53.51
Extended hypertensive	59.60
Infusion, DIP	**70,00**
Skull X-ray:	
Limited series	21.71
Complete (minimum of four views)	41.53

NA = not available

SOURCE Urban Institute sample of approximately 5.0130 solopract It loners, in cluding 177 radiologists (326 Percent of the solo radiologists in Call fornia)

Table I.— Estimated Number of Diagnostic X-Ray procedures in the United States, 1970 (in thousands)

Type of examination	All	Hospitals	Private office		Private groups	Health agencies and others	Average number of films per examination (all sources)
			Radiologists	Others			
All radiologic procedures	129,070	81,688	3,334	20,419	8,923	14,708	2,4
Skull	**4,220**	3,616	NA	NA	NA	NA	4.03
Chest radiograph	48,569	32,491	NA	7,565	3,207	4,293	1.7
Barium enema	3,428	2,774	NA	NA	NA	NA	**3.!34**
Excretory urogram	3,996	3,413	NA	NA	NA	NA	5.31

NA = not available

SOURCE Department of Health, Education and Welfare PublicHealth Service. *Population* Exposure *to* X *Rays,* U S 7970 November 1973 (30)

medical profession. These debates have virtually always been brought about by the findings of evaluative research. Indeed, evaluations of the benefits of these procedures seem to have raised rather than answered questions about proper indications for their use, especially when the results have contradicted prevailing patterns of medical care. Thus, a critical review of the evaluative research on each of the four diagnostic technologies is a good way to characterize the current controversies surrounding each of these procedures.

COST-BENEFIT AND COST-EFFECTIVENESS ANALYSIS IN THE EVALUATION OF DIAGNOSTIC X-RAY PROCEDURES

Cost-benefit analysis (CBA) and cost-effectiveness analysis (CEA) are methods to assist in allocating scarce resources among alternative uses. These methods were developed primarily to evaluate large public-sector investments such as highways, dams, and airports. When applied to diagnostic procedures, they are intended to provide information on two related questions: 1) Under what circumstances should the procedure be performed? and 2) How much investment in capacity to perform the procedure is justified? The answer to the second question rests on thorough study of the first, for only by knowing when a procedure should be performed can one assess how much investment in capacity is justified.

Applying the principles to medical procedures, CBA would enumerate and place a value on all benefits (both positive and negative) derived from performing a procedure on patients with a specified set of conditions and would compare those benefits to the cost of performing the procedure. The resulting net social benefit would indicate whether the procedure should be performed under the specified conditions. In traditional CEA, a measure of procedure effectiveness would be designated, and the ratio of that single measure to cost would be the critical item for resource allocation. Lives saved, life-years saved, quality-adjusted life-years saved, disability saved, and age-adjusted disability days saved are measures of effectiveness often chosen in studies of health care programs.

One can generalize the notion of economic evaluation to a social accounting framework, in which all dimensions of effect as well as cost are identified and their values estimated. It is difficult to argue with the soundness of knowing the direction and magnitude of all effects resulting from the performance of a procedure on a particular patient. However, the ideal of comprehensive and accurate evaluation is seldom met and may not be worth its own costs. Virtually all good evaluative studies are limited in the dimensions of cost and/or effectiveness under investigation.

The critical weakness of most studies reviewed in this paper is that they fail to consider one or more important implications of the procedure under study. By not dealing with these important dimensions of effect or cost, they leave open the possibility that their conclusions will be criticized or, worse, ignored. Yet, the paradox is that the conditions necessary to produce accurate information on the full array of effects and costs may not be achievable. The methodological and ethical problems of evaluative research, as well as its costs, frequently are barriers that cannot and perhaps even should not be overcome.

ORGANIZATION OF THIS BACKGROUND PAPER

This background paper is divided into two parts. The purpose of the first part is to summarize the different evaluative models underlying studies of the four diagnostic X-ray procedures and to lay out the strengths and weaknesses of each method. That part also identifies

the conditions under which these models are likely to provide information that can affect patterns of medical care. The second part of the paper contains four separate chapters summarizing what is known about the utilization, costs, risks, and benefits of each procedure, with particular emphasis on the evaluative methods employed.

The review of the evaluative literature of the past decade is by no means comprehensive. To be included here, a study must have provided evidence pertaining to at least one ot the following two interrelated questions. Under what conditions, if any, should a particular procedure be performed? And, how should the procedure be performed? Though many clinical studies address these kinds of issues, further conditions were required of those included in the review: measures of benefit, risk, or cost had to be specified explicitly; two or more alternative diagnostic strategies had to be compared either explicitly or implicitly; and a large enough number of cases had to be analyzed to draw meaningful inferences. The application of these additional filters eliminated editorial opinions and case reports from the review.

SUMMARY OF FINDINGS

This study has provided insight into three important questions. What influence has the evaluation literature had on the use of each of the four diagnostic X-ray procedures? What factors limit the influence of evaluative findings on medical practice? And, what directions might evaluative research take to increase its influence over medical decisionrnaking?

We cannot be precise about the influence of the evaluative literature on medical practice, because data are unavailable on rates of use of X-rays over time. But it is possible to infer from the clinical literature whether a consensus has developed in response to evaluative findings. The influence of evaluative studies of X-rays in screening (symptomatic) contexts appears to be strong, but evaluations of X-rays in diagnostic contexts, where patients present with complaints or symptoms, seem to have little impact on medical standards or practice. In either context, the more dramatic the results of the evaluation, the more likely is the study to have an impact. For example, studies demonstrating very low diagnostic yield of skull X-rays in emergency rooms created general concern and have led to some change in practice in a few centers (99,126).

The reasons for the limited influence of evaluations lie partly in the evaluative studies themselves and partly in the health care system. Study methods often are so flawed that the results cannot be trusted. Patient selection bias due to uncontrolled study designs are a major problem. Radiologic methods are often unstandardized. All too often, a procedure is evaluated in a group of patients so heterogeneous in its presenting signs, symptoms, and risk factors that the results offer no guidance at all about who should be X-rayed and who should not. The evaluative criteria rarely include the ultimate benefits of the procedure. Time and time again, as studies are reviewed in subsequent chapters of this background paper, we conclude that the study findings are inadequate because the implications of the X-ray results for patient health and well-being and for medical costs are unknown.

A more fundamental barrier to the use of evaluation in decisions about X-rays lies in the conflict between the individual patient's best interest and society's best interest. The individual patient, who seldom has to pay the full cost of the procedure and often need not pay at all owing to the availability of insurance or other third-party payment, need not consider the costs against the benefits of the examination. The public as a whole, however, must make these tradeoffs. This may explain why the results of evaluations of X-ray screening programs are more influential than are evaluations of diagnostic uses of X-ray. Screening programs are often funded by public health agencies, not by insurance coverage.

How can evaluations have a greater impact? In the diagnostic context, investigators can begin to consider the effect of the diagnostic strategy on therapy and ultimately on patient outcomes. Attention should be given to the clinical significance of missing disease if an X-ray is not ordered. Independently, research by economists and social scientists on the reassurance value of negative X-rays might put these benefits into proper perspective. Greater care might be taken in separating evaluations for patient groups with different presenting conditions. The appropriate level of aggregation of patient characteristics should be considered explicitly in the study design. To this end, studies of the diagnostic yield of X-ray and symptoms will be suggestive of patient groups where more thorough, outcome-oriented analysis is warranted.

Still, these new research directions will have no effect on decisions about the use of X-rays until both physicians and patients accept as reasonable the possibility that the diagnostic process might be truncated when the costs of pursuing additional information outweigh its potential benefits.